

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 29, 2002, 11:38:48 ; Search time 21.33 Seconds
(without alignments)
2914.665 Million cell updates/sec

Title: US-08-485-355B-50
Perfect score: 3374
Sequence: 1 MGDAGVASQRPNRCGRNV.....GKIARVARRARRARRAN 647

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2179	64.6	644	A43370	capsid protein - N
2	130	3.9	597	B75556	hypothetical prote
3	130	3.9	708	QOBBB8	UL80 protein - hum
4	126	3.7	3570	T45025	mucin MUC5B, trach
5	125.5	3.7	1269	A90267	proteinase related
6	125	3.7	2314	T28698	hypothetical prote
7	123.5	3.7	1204	C75015	glycoprotein 350/2
8	123	3.6	886	S29605	adhesin homolog 1m
9	122	3.6	1582	AC1153	membrane antigen g
10	121.5	3.6	907	QOBBE21	penicillin-binding
11	120	3.6	873	B75514	proteocadherin 68 -
12	119.5	3.5	889	T05055	hypothetical prote
13	119.5	3.5	2271	F90073	insulin-like growt
14	119	3.5	1363	T43220	probable peptidogl
15	119	3.5	1993	AF1450	G-cadherin - sea u
16	119	3.5	2809	T30213	hypothetical prote
17	118.5	3.5	3029	S76109	hypothetical prote
18	117.5	3.5	609	S62518	conserved hypothet
19	117.5	3.5	884	E75489	titin, cardiac mus
20	117.5	3.5	26926	I18344	hypothetical prote
21	117	3.5	674	A97490	hypothetical prote
22	117	3.5	859	AE2217	hypothetical prote
23	116	3.4	3716	E70969	chitinase PPE prote
24	115.5	3.4	699	A38368	large repetitive p
25	115.5	3.4	3624	AD0835	agglutinin-like ad
26	115	3.4	1419	T30531	membrane antigen g
27	114.5	3.4	710	QOBBE22	probable transcrip
28	114	3.4	828	T12184	coat protein VP1 -
29	113.5	3.4	781	VCPV19	

30	113	3.3	4436	2	E71086	hypothetical prote
31	112.5	3.3	604	2	A84753	hypothetical prote
32	112	3.3	1090	2	S59077	cellulose 1,4-beta
33	112	3.3	1578	2	AD1512	peptidoglycan boum
34	111.5	3.3	604	2	T31042	hypothetical prote
35	111	3.3	1902	2	B45764	lactosepin (EC 3.4
36	110.5	3.3	505	2	S39562	endoglucanase - Er
37	110.5	3.3	677	2	E75563	probable peptidyl-
38	109.5	3.2	792	2	F88656	protein F56D6.1 li
39	109.5	3.2	1662	2	T18540	mofta protein precu
40	109	3.2	591	2	B87361	flagellar hook pro
41	109	3.2	766	2	B86597	CT711 hypothetical
42	109	3.2	766	2	F72027	conserved hypothet
43	109	3.2	1302	2	T30191	multidrug resistan
44	108.5	3.2	642	2	B72428	laminarinase - The
45	108	3.2	424	2	T32434	hypothetical prote

ALIGNMENTS

RESULT	1
A43370	capsid protein - Nudaurelia omega virus
C:Species: Nudaurelia omega virus	
C:Date: 16-Jul-1999	#sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: A43370	
R:Agarwal, D. K.; Johnson, J. E.	
Virology 190, 806-814, 1992	
A:Title: Sequence and analysis of the capsid protein of Nudaurelia capensis omega vir	
A:Reference number: A43370; MUID:92391097	
A:Accession: A43370	
A:Status: preliminary	
A:Molecule type: mRNA	
A:Residues: 1-644 <AGR>	
A:Cross-References: GB:S43937	
C:Superfamily: Nudaurelia omega virus capsid protein	
Query Match	64.6%; Score 2179; DB 2; Length 644;
Best Local Similarity	64.7%; Pred. No. 2e-145;
Matches 425; Conservative 68; Mismatches 130; Indels 34; Gaps 7;	
QY	8 SQRPNRGRTRNV-----SANTVYNGRRNRGRGRQVSPDNFTAAQDLA 57
DB	3 SNSASGRSRNRRIARQARRARRSANVYNAFK-----QRANNIDNTAAQDLG 54
QY	58 QSIDANTVTPPANISSMPEFRNNAKGIIDSDSIGWTKYIDPAGATESARAVEYSKI 117
DB	55 QSIDANTVTPPANISSMPEFRNNAKGIIDSDSIGWTKYIDPAGATESARAVEYSKI 114
QY	118 POSLVKFSVDAETREIYNEDCPVYTDVSVPLDGRQSLSFSPMERITVYVANENKE 177
DB	115 POSLVKFSVDAETREIYNEDCPVYTDVSVPLDGRQSLSFSPMERITVYVANENKE 174
QY	178 MSIDVNVNDLIEVNLNADRWYVDSQWINTDTYVYRIYRLRPTVDVPPTEGLVLT 237
DB	175 ISLDVNDLIVMNLNADRWYVDSQWINTDTYVYRIYRLRPTVDVPPTEGLVLT 234
QY	238 VSDYRLTYKAITCEANMPTLVDOGFVIGQVYALPTSLPQYDVSEAYALHTLTFARPSSA 297
DB	235 CSQYRLTYKAITCEANMPTLVDOGFVIGQVYALPTSLPQYDVSEAYALHTLTFARPSSA 294
QY	298 AALAEVWAGLPQGGTAPAGTAPAMEQSSGGYLTWRHNGTTPAGSVSYVLPFGPALERD 357
DB	295 AGVTLTWASMPGGSAPSDPAIIPDSTIQF-QMRHGGFDAPPGVYTYITIPAGTYWQIFD 353
QY	358 PDGSGWTFASAGDTVFQVAVDEVVYVNNPAGCG-----SAPFTTVV-PPSNAYTNT 411
DB	354 TTNENNGFANPDVVFQ-----TGAGAGTATITITAPVTLTILTTTSAANV 405
QY	412 V-FRNTLTETROSSRRLLEPMPEADFGQTVANNPKTEGSLKETIGCYLVHSMKRPVFO 470
DB	412 V-FRNTLTETROSSRRLLEPMPEADFGQTVANNPKTEGSLKETIGCYLVHSMKRPVFO 470

Db 480 EPPSGEMRYDGHGRGOSKOHRRHSGSGCHNKRKRTAAASSSSDDLSFPGFAEHRGAR 539
 QY 534 ---VTWVNPFGFAGAGLLKNEILLADDLATRLTGVPADNNAVASAFAMMSS 590
 Db 540 KRKSIVNNDGGSGGAG--SNQOQORDELDALHEI--KRDLFAAQS--STLLSA 592
 QY 591 VLKSEATSSIIKSVGETAVGAAGSLAKPLGLIMSPVKIAARVA 636
 Db 593 ALPSAASSPTTTTCTPTGELTSGGGERTFALLSGCAVAERAAQ 638

RESULT 4

T45025

muscn MUC5B, tracheobronchial [imported] - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000

T45025

R:Desseyn, J.L.; Guyonnet-Duperrat, V.; Porchet, N.; Aubert, J.P.; Laine, A.

J. Biol. Chem. 272, 3168-3178, 1997

A:Title: Human mucin gene MUC5B, the 10.7 kb large central exon encodes various alternat

A:Accession: T45025

A:Reference number: 222899; MUID:97166151

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3570 <DES>

A:Cross-references: EMBL:72496; NID:g1834502; PIDN:CAA96577.1; PID:g1834503

A:Experimental source: placenta

C:Genetics:

A:Gene: MUC5B

Query Match 3.7%; Score 126; DB 2; Length 3570;

Best Local Similarity 22.3%; Pred. No. 11;

Matches 94; Conservative 32; Mismatches 178; Indels 118; Gaps 14;

QY 221 LRPTVDPPTGELVTRVSDYRLTY---KATCEANMPLVDGFWIGGQYALPTSLP 276
 Db 643 LRSTATTPATSTPTAIPSSILGTTWRLSQTTTPMATST-----ATPSSNP 689
 QY 277 QYVNSAVYALHTLT-----FARPSAAALAFVWAGLPQGGTAPAGCPAMQ 322
 Db 690 E-----TVHTSTVLTATTTGATGSAVATPSSPTGTAHTKVLTTTGTATPSSSP 742
 QY 323 ASSGGYLTWRHNGTTPAGSVSYLPEGFALERYDPNDGSMWDFASAGDTVFEROVADE 382
 Db 743 GARATLPVW-ISTTTPTTRGSIPTPSST-----PGTHHTPTVLTITT 784
 QY 383 VVVTNNPAGGSAPTFTVVPNSNATNTVFERNLTLETRPSRRLELPMPPADFGQTVAN 442
 Db 785 TVVATGSMATPSSSTQTSPTSLTTATTTATGCTNPSSPTGTPPIPVV--LTTTAT 842
 QY 443 NPRIEISLKEITLCVLIYSKMNPFQOLPPASSGCAVFPNNCGYERTDLPDYGIRIS 502
 Db 843 TPAATSS-----TVTPSSALGTT-----HTPPVNP----- 867
 QY 503 FQDNMTAAVAFHRSLSHS-----CSIYTKYOGMEGVTVNTPFGOFAN-----AGLKN 552
 Db 868 -----TTATTHGRSLSPSPRHVCTAMTATGATGILGTHHTERTSHTPRAITGTQH 922
 QY 553 EEILCIADDLATRLTGVPYR---TDNFAAVSAFAA-----NMSSVLKSEATSSIIK 602
 Db 923 STPALSPHSSRTSPSPGTTTGTGHTTATSRITATAPSKTRSTLLPSQPTSAPI 982
 QY 603 SV 604
 Db 983 TV 984

RESULT 5

A90267

proteolase related protein [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001

C:Accession: A90267

R:She, O.; Siroch, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyer, M.J.; Ch

Jong, T.; Jeffries, M.A.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder

aretz, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: A90267

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1269 <KOR>

A:Cross-references: GB:AEO06641; NID:g13814328; PIDN:AAK41392.1; GSPDB:GN00155

C:Genetics:

A:Gene: SSO1141

Query Match 3.7%; Score 125.5; DB 2; Length 1269;

Best Local Similarity 19.6%; Pred. No. 2.4;

Matches 98; Conservative 69; Mismatches 181; Indels 153; Gaps 27;

QY 18 RNVRSANVTYNGRRNRRTGRQVSPPDNFTAAQDLAQSOLDANTVTFPANISSMPEF 77
 Db 584 ORLVNTANTYPNG-----SPVQ--TGEFKALIENTLGNLTFTNLTSLTKL 629
 QY 78 RNMAGKIDLSDSIGWYFKYLDPAAGATESARAVEYSKIPGLVKSVDAREITY--- 134
 Db 630 --WT--GSGVLSKSKASGILFVYV--YSSDGLKIGIGYEFESGYITFTNTTFFVVEL 684
 QY 135 -NEECPVTVDS--VPLDGRQWSLISFSPMFRAYAVAN--VENKESLDVNDL 186
 Db 685 GNAELGITLSNSYFQAPIDVMNITLNIYSNTTNAVTFVTLVPIKNGVVIDLPDL 744
 QY 187 IEMNLAMRVVNDSEQ--INFMDT--TYVYRIY-----LRP 223
 Db 745 -----SIGDLITAEENAGFDAPFTNGVYMTFLIPQVYVPGSVSPQHTIGSTIP 799
 QY 224 TYDVPDPT--EGLVTRVSDYRLTYKAITC-----EANNPTLVDOGFWIGGQYALPTSL 275
 Db 800 PYNLPSTTFQDALQGT-----NITAKLVSSNGVINEANIP--LSPNGIYRG--YLVIPKMT 852
 QY 276 PQ--YDV-----SEAYLHLTLTPARPSAAL-----AFWAGLPQ 309
 Db 853 PSGLVNVLLFAVYSYTLNTLTIRGFYQIYVSNQATISVKSVAFAEGQTFIVANITIN 912
 QY 310 GGA-----PAG-----TPAMQASGGGYLIMRHNGT--TFPAGS 342
 Db 913 GINEIRFGMFSATVYIPSSLSFNITTISSIIIEIPLMNPKIG--EHEGNFTLPSAISGN 969
 QY 343 VSVYLPEGFALERYDPNDGSMWDFASAGDTVFEROVADEVVVTNNPAGGSAPTFTV 402
 Db 970 LTYLAGQYF-----GVPEK-VLIGISALGNPTTNGSNMYTIV 1009
 QY 403 PPSNATNTVFERNLTLETRPS 423
 Db 1010 LPYTLFTN-----QTLDKTLPS 1026

RESULT 6

T28698

hypothetical protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T28698

R:Parkhill, J.; Bentley, S.D.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, January 1999

A:Reference number: Z20512

A:Accession: T28698

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2314 <PAR>

A:Cross-references: EMBL:AL023496; NID:e1292348; PID:e1292365; PIDN:CAA18915.1

Query Match	3.6%	Score 123	DB 2	Length 886
Best Local Similarity	20.38%	Pred. No. 2		
Matches 138	Conservative 66	Mismatches 269	Indels 206	Gaps 29

RESULT 9
AC1153
adhesin homolog lmo0627 [imported] - *Listeria monocytogenes* (strain EGD-e)
C:Species: *Listeria monocytogenes*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AC1153
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Dutand, L.; Dussurgeir, O.; Entlian, K.D.; Fsihi, H.

Query Match	3.6%	Score 122;	DB 2;	Length 1582;
Best Local Similarity	19.68;	Pred. No. 5.9;		
Matches 122; Conservative	92;	Mismatches 247;	Indels 160;	Gaps 35

RESULT 10
Q0BE21
membrane antigen gp350 - human herpesvirus 4 (strain B95-8)
C.Species: human herpesvirus 4, Epstein-Barr virus
C.Date: 25-Feb-1995 #sequence,revision 25-Feb-1995 #text_change 28-Jul-2000
C.Accession: A43042; S33008; S33009; A03762
R.Bankier, A.T.; Delninger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983

A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
 A:Reference number: A93065; MUID:85035713
 A:Accession: A43042
 A:Molecule type: DNA
 A:Residues: 1-907 <RAN>
 A:Cross-references: GB:V01555; GB:J02070; GB:K01729; GB:K01730; GB:V01554; GB:X00498; GE:R:Farrell, P.J.
 submitted to the EMBL Data Library, March 1988
 A:Reference number: S32973
 A:Accession: S33008
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-907 <PAR>
 A:Cross-references: EMBL:V01555; NID:959074; PIDN:CA24855.1; PID:91334869
 A:Accession: S33009
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-907 <PA2>
 A:Cross-references: EMBL:V01555; NID:959074; PIDN:CA24855.1; PID:91334869
 R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H
 Nature 310, 207-211, 1984
 A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
 A:Reference number: A03794; MUID:84270667
 A:Contents: annotation: protein coding region
 C:Superfamily: Epstein-Barr virus membrane antigen gp350

Query Match 3.6%; Score 121.5; DB 1; Length 907;
 Best Local Similarity 19.6%; Pred. No. 2.7;
 Matches 138; Conservative 66; Mismatches 266; Indels 233; Gaps 28;

QY 48 NPTAAADLAQSLDANTVTPPANISSMPEFRNNAKGIKIDDSIGYFYFLDPAGATES 107
 DB 169 NITAVVR--AQGLD---VTLPLSLPTSADSNS---VKTEMLG---NEID---IEC 211
 QY 108 ARVAGEKSLPDGLVKSVAEIRE-----IYNECPVTVDSVPLDGRMSISFSEF 160
 DB 212 IMEDGELSQLPGDNKNFNTICSGESHVPSGGLTSTSPVAT--PIGCTGATSLRLTPR 269
 QY 161 PMFR-----TAYVAANVENKESLD-----VNDLI-----EMLNLADMRYVDSE 203
 DB 270 PVSRLFGNNSILVYFSGNGPKASGDYCIQSNIVFSDDELPSAGDMPTNTDTIYVGDN- 328
 QY 204 QMNFNTDTTYVIRIRLPTTYDPTDEGLVTVSDYRLTYKAITCEANMPLLVDOGEF 263
 DB 329 -----ATYSVPMTSE-----DANSPVTVTAFF 352
 QY 264 -----IGGOYALPT----- 273
 DB 353 AMRNNETDKCKWTLTSGPSCGENISGAFASNRFTDITVSGLTGTAAPKTLITRTATNA 412
 QY 274 -----SLPOYDSEAYALHTLTFARPSSAALAFAVAGLPGOGTAPAGPAMEQAS 324
 DB 413 TTTTHKVIKAPESSTTSTPLNTGTGADPNTT-----TGLPSSHVPTNLTA--PAS 463
 QY 325 SGGYLTRHNGITFPAGSVSYLPEGFALERYPDNGSW-----TDFASAGDIY 373
 DB 464 TGPTVSADVTSPPTAGTSGASP-----VTPSPSPMNGTESKAPDMTSSSTPTTP 516
 QY 374 TFRQVAVDEVVVT-----NNPAGGSAPTFETVAVPPSNAYNTNFR 414
 DB 517 TPNATSTPRAVTTPTPNATSTPRAVTTPTPNATSPITGKTSPTSAVTTPTPNATSPITLK 576
 QY 415 N-----TLLETRPSSRLLELPPPADFGOTVANNPKIEOSLLEKTLGKY 458
 DB 577 TSPTSATVTPTPNATSPITLKTSPTS-AVTTTPPNAT--GTVGETSP--QANNATNHTLG-- 631
 QY 459 LVHSKMRNPVQLTPASSFEAVSPNNNGYERTDLDPYTIKOSFDONMSTAAVHFRSLS 518
 DB 632 ---GTSPTPVVTSQPKNATSAVTTGQ--HNITSSSTSSMSLRPSSNRETLSPSTSDNSTS 686
 QY 519 HSCSIYTKTYGMEGVNTVTPFGGFHAGLLANNEILCLADLATPLTGVVAT----- 573

DB 687 HMLPLTSAPHTGENTITOV--TPASISTH-----HVSTSSPAPRGTTSSQAS 731
 QY 574 --DNFAAVSAPFANMLSVLKSEATSIKISVGETAVGAAQS 614
 DB 732 GPGNNSSTSTPKEGVNTYKGTTPPNATSPQAPSGCKATVPTVS 774

RESULT 11
 B75514
 penicillin-binding protein 1 - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: B75514
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Yamachavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: B75514
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-873 <WHI>
 A:Cross-references: GB:AE001907; GB:AE000513; NID:96458162; PIDN:AAF10059.1; PID:9645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR0479
 A:Map position: 1

Query Match 3.6%; Score 120; DB 2; Length 873;
 Best Local Similarity 26.0%; Pred. No. 3.2;
 Matches 72; Conservative 19; Mismatches 104; Indels 82; Gaps 16;

QY 208 FTNDTTYVYRIRVLR-----PTYDVDPPT-----EGLVTVSDYRL 243
 DB 545 FANGGLYAPSAVRHEDPOGKVLKRPDPVGRWMDKRYAMGLDMIRGVNDLSAYQ- 603
 QY 244 TYKAITCEANMPLLVDOGEWIGGOYALTPTSILPOYDSEAYALHTLTF--RPSAALA 301
 DB 604 --GGLATFRARI-----DGMQVGGKGTGT-----NDVKD-----LWFAGVPTLTGSA-- 642
 QY 302 FYVAGLPGGTAPAR-----GTPAMEQASSGGYLWRHNGTTFPAGSV-----SYLPLPG 350
 DB 643 -VWVGKOEGETLPSMAVSGEVPPTPVW-QAATGAGLAGRETATFIPGGLVYRTYRLMA 700
 QY 351 FALERYP-----NDGSWTFEASAGDTVTFROVAVDE--VVVTNNPAGGSAP-----T 397
 DB 701 FLEENMDQDVRHDS-----RRATXXATPRAVYAPERPQPATITTPAGRSAPLRRTTRT 756
 QY 398 FTVRVPPSNAYTNTVFRNTLLETTPSSRLLELPPPA 434
 DB 757 LPRQTSFRRLPS-----RTCRPRRPLSPPPRRRPGA 789

RESULT 12
 T09055
 protocadherin 68 - human
 C:Species: *Homo sapiens* (man)
 C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
 C:Accession: T09055
 R:Jin, P.; Xu, H.; Israel, D.
 submitted to the EMBL Data Library, October 1997.
 A:Reference number: Z16540
 A:Accession: T09055
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-889 <JIN>
 A:Cross-references: EMBL:AF029343; NID:92599501; PID:92599502
 C:Genetics:
 A:Gene: FCH68

us-08-485-355b-50.rpr

Query Match Summary 3.5%: Score 119.5; DB 2; Length 889; Best Local Similarity 20.4%; Pred. No.3.6; Matches 144; Conservative 92; Mismatches 230; Indels 219; Gaps 36;

```

32 RNRORRTGROYSPPDNTAAADLAQSLDANTVTPFANISSMPEFRMWAG----- 83
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
97 RHNAKCOLSEVFANDKEICKMIKVEIODINDNAPFSSDOIIXLIDSENAAGTRPEPLSA 156
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
84 -KIDDSISIGYFYLLDPAGATE-SARAVGEYSKIPDGLKFSVDAIRETYNECPYV 141
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
157 HDPDGEAGGLRYLLTRRDHGLFGLDVKS RGDGTFFPLVLOKALD---RDOQHHTLVL 213
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
142 TDVSVPLRGROWSISIFSEPMFRTAYVAVANENKENS--LDVVDLTETMLNNLADMRV 199
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
214 T-----ADLGG-----PPRSATQIVNKVYIDSDNNSPVEAPSYLVELENTPLGIV 262
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
200 VDSEQWINFN-----DTTY---VYRIRVLPFTYDVPDPEGLVTVS--DY----- 241
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
263 ID---LNATDADEGPNGEVLVSFSSYVPRDY-RELFSI-DPKTGLIRVKCNLDEENGM 316
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
317 LEIDVQARDLGNLIPAKCKYTKLIDRNDNAPSI---GF-----VSRGALISE 363
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
283 AYALHTLTFAPRSSAALAFW--AG-----LPQGTAPAGTPAMBOASSGGYLT 330
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
364 A-----APGCTYIALRVTRVDRSGKNQLOCRVLGGGTGGG----- 401
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
331 WRHNGTTPAGSVSVLDEGFALERYDPDNGSDWDFASGDTYPRQY---AVDEVVYTN 387
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
402 ---GLGGGGSVPRKLENT-----DNFT-----VVDPRPDRETQDEYNTI 442
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
388 NPAGGSAF---TFYRV---PE---SNATNTVFENLTLETSPSSRRLLELMP 433
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
443 VARDGSPPLNSTRKSFALKILDKNPNPRFTKGLVLOVHENNI-----PG 488
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
434 ADFOGTVANNRKIEQ---SLKETLGCYLVHKKM-NRYFOLTTPSSSGANVSFNNPG 486
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
489 DYLGSVLADODPDLGNGTGVSYSLIPSHIGDVSITYVSVNP---TNGAITYALRSFN-- 541
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
487 YERTSDLDPTYTIRDS-----FDQNSTAVAHFRSLSHSCSITVFTYGGWG----- 533
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
542 EQGTAFERKVLAKOSGAHLESMATYRVTVLDVNDNAPYIVLPTLONDRAELQVPRNA 601
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
534 ----VTNVNTPFGCFPAHAGLKK-----NEELCLADDLATRLTGVPATDN-----F 576
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
602 GIGYVSTVRALSDPFGESGRLTVEIVDGNDDHLFEIPSSGEIITLHPFEDVYVVEL 661
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
577 AAASAPAAANMLSSYLKSEATSSIIKSGTFAVGAOSGLAKLP 621
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |
662 VKVTVDHCKPTLTSAVAK-----LITRSVS---GSLPEGVPRVNG 697
   | : : : : : : : : : : : : : : : : : : : : : : : : : : |

```

RESULT 13

F90073

biochemical protein SA2447 [imported] - Staphylococcus aureus (strain N315)

C.Species: Staphylococcus aureus

C.Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C.Accession: F90073

R.Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Iji, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hirata, S.; K. Lancet 357, 1235-1240, 2001

A.Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A.Reference numbers: A89758; MUID:21311952; PMID:11418146

A.Accession: F90073

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-2271 <KUR>

A.Cross-references: GB:BA000018; PTD:913702612; PIDN:BBB43752.1; GSTDB:GN00149

A.Experimental source: Strain N315

CitGenetics:

A:Gene: SA2447

Query Match similarity 3.5%; score 119.5; DB 2; length 2271;
Best Local similarity 18.9%; Pred. No. 15;
Matches 147; Conservative 107; Mismatches 233; Indels 289; Gaps 37;

OY 22 VSANVTVN-----GRNRQRRRTGROVSPEDNFTAAADLAOSLDANTVPPEAN 70
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 254 VTANITIVTKNDLAKOYMTTSGNATYDOSTGTIVTLQDAYS-----QKGAITLTGR 303
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 71 ISSMEEFNNMAKKRIDDS-----DSICWYRK-----97
 ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 304 IDSNSKFH--PSGKNVLANKYEGHGNGGDGIGFAFSPGLGETGLNGAAGIGLSNAFG 361
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 98 ----YLDPAGATESARAVGEYSKI-----PDGLVKRSV 126
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 362 FKLDIYHNHTSKPNSKAANKADPNANAGGACGAPITTSYGVAATTYTSSSTADNAKINAY 421
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 127 -----DAETREIYNEECPVYTDVSVPLDGROWSLSI-----RSPFM----162
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 422 OPTNTPFODFDIN--YNGDDTKVMT---VKYAGQTWRNISPMIAKSGTNFSLSMTASTG 476
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 153 -----FR-AYVAANVANENKEMSLDYVNDLI--EMLNLMDMRVRVNSEO-----204
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 477 CATNIQQVOQGFETEESAVTVQVRVDYTGKDIIIPPTTYSGNQDVYTTIDQOSALTAK 536
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 205 WINFTN-DTTYVYIRLVRLPYVDPEPTEGLVRYVS DYRLTYKAITCEANNPTLVDGFW 263
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 537 GNNTSYSDSSY-----ASTYN--DTNKTKVMTNAGGSVITYFTFDKAPVTVGNQIE 587
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 264 IGGGVALPFTSLPOVDSEALHLHTLFARPSSAAALFVAAGLPQG-----GARPGT 317
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 568 VG--KTMPNPVYLLTTDNGTGVNTVT-----GLPSGLSYDATNSITIGT 630
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 318 P-----AMEOASSGGYLTWRHN--GTFEPAGSYYLYPECG-PALERYPD-----358
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 631 PKIKGSIVTVYVSTDQAANKSTTFETINVDTLAT-----VTPICDQSSEVYSPIRIK 685
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 359 -----NDGSWTDFASAGDTVFTRQVADEVVYVNNPAGGSAFPF-----TVRVPPSNAY 408
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 686 IATQDNSEN-----AVTNVTVTLPSGLTFPDSYNNITISGTPINIG 724
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 409 TNT-----VERNTLLETRPSSRLLELMPPADGGOVYANPKIEQ 448
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 725 TSTIIVYSTDASGNKTTTTFKEVTRNSMSDVSSTGSEQ-----SQSVTSKADSQ 777
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 449 SLKETLTCYLVHSKMRNPVQLTPASSFGAVSFEN-NGYERTRDLPYTGIRDSFDQM 507
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 778 SASSTSSGIYS-----TSASTKSSTSVSLSDSVSASKLS--TESNSNVSSST 825
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 508 STAAHFPSLSHSCS-IYTKTYOGMEGVYVNTVPPGOFHAAGLKNEI-----555
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 826 STLINVSQSVSSSMGYSKSTSLSDSISNSNS-----TEKSESISTSTSDELRTS 876
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
OY 556 LCLADDLTRLTGVYPATDNFAAFSAFAANMLSVLSEATSSLIK---SVGEDA 608
 |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 877 TLSLDSLMSSTSGSLSKSQSLSTISGSSS---TSASLSDSTSNASTSTLSESA 929

RESULT 14
T43220
Insulin-like growth factor-1 receptor - common lanelet
N.Alternate names: insulin-like peptide receptor
C.Species: Branchiostoma lanceolatum (common lancelet)
C.Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-May-2000
C.Accession: T43220
R.Pashmforoush, M.; Chan, S.J.; Steiner, D.F.
Mol. Endocrinol. 10, 857-866, 1996
A.Title: Structure and expression of the insulin-like peptide receptor from amphioxus
A.Reference number: Z22346; MUID:96408719
A.Accession: T43220
A.Status: preliminary; translated from GB/EMBL/DDBJ
Molecule type: mRNA

A:Residues: 1-1363 <PAS>
 A:Cross-references: EMBL:SB83394; NID:g1911171; PID:g1911172; PIDN:AAB50848.1
 C:Superfamily: Insulin receptor; protein kinase homology
 C:Keywords: hormone receptor

Query Match 3.5%; Score 119; DB 2; Length 1363;
 Best Local Similarity 19.8%; Pred. No. 7.6;
 Matches 99; Conservative 60; Mismatches 162; Indels 180; Gaps 23;

QY 79 NMAAGKIDLDSDSIGWYK-----YLDPAATESARAVEYSKIPDGLKFSVDA 128
 Db 416 DMDTDITIDGKLFHFHNPRLCHVLTWADKGLPEHA-----ITD----- 458
 QY 129 EIREIYNEECVPVTVDSVPLDGRQMSLIFSFPMEFRTAYAAVANENKMSLDVNDLIE 188
 Db 459 -----TDISTLTFNGDQ-----AQCSFSRL-----IE 480
 QY 189 WLNLADMRVYVDEQWINFNDTTYVRIYRLVLPDYDVPDPTGGLVRYSDRLTYKAI 248
 Db 481 EINTSKMILL---RMSEF-----RP---DPPRLLSYTVS-YRET--- 514
 QY 249 TCEANMPTLVDOGF-WIGQVYALPTSLPOYVSEAYALHTLTFARPSSAAL---AFVW 304
 Db 515 -----EDQIGIDYDGDQDAGCNTEMKEPDSPTQTALHTGLKMTQYALLVKITYK 565
 QY 305 AGLPQGG-----TAPAGTPAMEQASSGGYLTWHRNGTTEPAGSVSYLP 348
 Db 566 AGAEGSGAKSDIYARTADKPTHPQDVVYSSNTLITMKPRNR---PNCNVTHYIV 623
 QY 349 EGFALEIYDNDGSMTFDASAG-----DYTFRQYAAVDEVVYNNPAGSGSAPFTYRVP 403
 Db 624 K-YKROEDVAEMEQRYCKGGLKPHRPTQGLIEDIVNNEEPNNSTIGDGTG---CECP 678
 QY 404 PS-----NAYNTYFRNTLTETRPSSRLLEPMPAEF---GQTY--- 440
 Db 679 KSEDEIRIEEBAAFQGEFENFLNNYHKKRENTRAGRRLRELPTVATRFYSQTYNTV 738
 QY 441 -----ANPKIEOSLKTLCGLYVHSMRN---PVFOLTPASSFGAVSFN 483
 Db 739 LPSTNRTVPTPTPNPPOLETTYPWNBH---VVLTLGRHFSEYIIEVICMADAAVGS 795
 QY 484 NPGYERTRLPDYTGIRDSFD 504
 Db 796 GSAVELARQAD-----DSAD 811

RESULT 15
 AF1450
 probable peptidoglycan bound protein (LPXTG motif) lin0141 [imported] - Listeria innocua
 C:Species: Listeria innocua
 C>Date: 27-Nov-2001 #sequence_rev1sion 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AF1450
 R:Laaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.; Jones, L.M.; Karst, U
 Science 294, 849-852, 2001
 A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Magueno, E.; Maltournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tjerez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MID:21537279; PMID:11679669
 A:Accession: AF1450
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1993 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC95374.1; PID:g16412560; GSPDB:GN00178
 A:Experimental source: strain Clp11262
 C:Genetics:
 A:Gene: lin0141

Query Match 3.5%; Score 119; DB 2; Length 1993;
 Best Local Similarity 19.3%; Pred. No. 14;

Matches 152; Conservative 100; Mismatches 276; Indels 260; Gaps 38;
 QY 3 DAGVASQRPHNRGRTRVRSAMTVYNGRRNRORRGROVSPDNDNTAAODLAOSLDA 62
 Db 1247 DASVA-----TAKTLIDGSVALVGGSTN-----GKTIKRTD-LITALKIDFTEKK 1291
 QY 63 NYVTEPANISSMPEFRNMWAKTI---DLSDSIGWFKYLDPAATESARAVEYSKIP 118
 Db 1292 STA-----PLYVTTANGKVAQKDELKIGANTLETKDAVVLSE-AAAVTK----- 1336
 QY 119 DGLVKSVDALIEIYNEECVPVTVDSVPLDGRQ-----WLSIFSFPMEFRTAYAAVAN 172
 Db 1337 ---AKTAAFEENKNGVNSDDEPRTD-VKVEQQLALKNGSRGGTYPL---TFLEKN 1389
 QY 173 VENKMSLD--VYNDLIEWNLNAD-----WRYVVDSEQWINFNDTTYVRIYRLV 222
 Db 1390 GKEVEYIDVKVEKDLLE-VNAHDSITYGDMWRADNDSALNKEGETLTFADLEARG 1447
 QY 223 P-----TYVDPDPTGGLVRYSDRLTYKAITCEANMPTLVDOGFWT----- 264
 Db 1448 TVDTTRAGEYPVYKYNDTKVTYTKD-----DATEINAHDSITYGDTWAKDNFDS 1502
 QY 265 -----GGQVYALPTSLPOYVSEAYALHTLTFARPSSAALAFVWAGLPQGGTAPAGTPA 319
 Db 1503 AADRQGENVALSKVYTVNTVNTAQAGTYPTTY-----YGGVSKITYTV 1546
 QY 320 WEQASSGGYLTWHRNGTTEPAGSVSYLPEGFALERYDNDGSMT---DFASA---GDT 372
 Db 1547 VKENKKG---INAHNMTIYV-----DSMTEDNFDAVVDKGNP 1583
 QY 373 VTRQYAAVDEVV-VTNPNAG-----GGSAPFTYRV---PSNAYNTYFRNTLLE 419
 Db 1584 VESKVTYVETPTVNTNKKAGTYOLKTYTDPGASVTVLTKYKNIQTAVANAHNSYVGETWE 1643
 QY 420 TR---PSSRLELPMPADGCTVA-NPKIEOSLKTLCGLYVSK-----M 464
 Db 1644 AKDNFDSARKD-----GETVAPADVBEQNVMTVAGTISTYTKYDGSKITIKVTY 1695
 QY 465 RNFVOLTTPASSF-----GAVSFNNGP--- 486
 Db 1696 KNPQTAITAHDSVYVYGDWMSAKDNFSAIDKAKRPVAYRDTVEEDPTVDINTPTGYSV 1755
 QY 487 ---YE-----RTRDLPDYG---IRDSFD-----NMSTAAVNH 513
 Db 1756 TYRYGGISKVVQITVKPRQTKYESHDSITYAGSMWAKDNFSAIDKKGDDVYKLSOVTVI 1815
 QY 514 FRLSHSCSIYTKYQWEGVTNVNTPFGQFAHAGLCKNEELCLADDLATRLTGVPAT 573
 Db 1816 GRVNDQTPGYEETIR-YGQVTSVSRVYVLQNHAKIIVDSKL-----KINMDWDAK 1866
 QY 574 DNFAAASAFANMLSSVLKSEATSSIIKSVGETAV-----GAOSGLAKL--PGIL 623
 Db 1867 DNFVRMSRDSSEIRPMSKYKVBGKVN-TKKAGKYQTYTIDPNEGIVDAGKRELSTVATI 1925
 QY 624 MSVPGKIA 631
 Db 1926 EYVEGKVA 1933

Search completed: May 29, 2002, 11:40:28
 Job time: 100 sec